Krister Bykvist¹

Person-affecting and non-identity²

According to a popular version of the person-affecting idea of morality, what is better (worse) must be better (worse) for someone. However, there seems to be a clear tension between this idea and some of our considered judgements about cases in which the existence of future people is contingent on our choice. For example, we want to say that creating a very unhappy person makes the world worse, other things being equal. In order to comply with a person-affecting morality in this case, we need to show that coming into existence can be worse for a person, but it does not seem plausible to say that it can be worse for a person to exist than not to exist. This paper discusses some recent attempts to ease this tension, and it is argued that none of these attempts is convincing. That leaves us with only one option: to reject the person-affecting constraint in its current form.

¹ Institute for Futures Studies & Department of Philosophy, Stockholm University, krister.bykvist@iffs.se.

² Financial support by Riksbankens Jubileumsfond (the Swedish Foundation for Humanities and Social Sciences) is gratefully acknowledged.

1. Introduction

One important part of morality is concerned with what is better or worse for people. According to a popular version of this *person-affecting* idea of morality, what is better (worse) must be *better (worse) for someone*.³ However, it is unclear how we are to put this idea to use in non-identity cases, i.e., cases where, depending on what we decide to do, different people will come to exist in the future. Indeed, there seems to be a clear tension between the person-affecting idea and some of our considered judgements about non-identity cases. In at least some non-identity cases we want to say that one outcome is better (or worse) than another in virtue of the wellbeing of people who do not exist in both. For example, we want to say that creating a very unhappy person makes the world worse, other things being equal. But how can we say this, if an outcome is worse only if it worse for someone? In order to comply with a person-affecting morality in this case, we need to show that coming into existence can be worse for a person. But can it really be worse for a person to exist than not to exist, and thus better for her not to exist than to exist? That seems to require that the person would have been better off not existing, which sounds paradoxical.

In this paper, I am going to discuss some recent attempts to ease this tension. According to these attempts, we can stick to a person-affecting morality and still avoid the counterintuitive judgement that no outcome is better or worse in virtue of the wellbeing of people whose existence is contingent on our choice. I shall show that none of these attempts is convincing. That leaves us with only one option: to reject the person-affecting constraint in its current form.

In section 2, I shall say more about non-identity cases, and list the most morally salient ones. In section 3, I shall make more precise what a person-affecting morality amounts to. In section 4, I shall present an argument that spells out the tension between person-affecting morality and our judgements about non-identity cases. The argument's conclusion is that no outcome can be better or worse than another in terms of the well-being of people who do not exist in both. In sections 5 to 10, I shall discuss possible ways to resist this argument while sticking to a person-affecting morality. I shall especially focus on the approach recently defended by the so-called 'Scandinavian existentialists'.⁴ I shall argue that the main problem with their approach is that they fail to fully acknowledge what it means to say that an abstract state of affairs has value.

³ See Temkin (1993a), (1993b), and Holtug (1996). The label "Person-Affecting Restriction" was introduced by Glover (1977), p. 66, but see also Narveson (1967).

⁴ See, for instance, Arrhenius & Rabinowicz (2010), (2015), Johansson (2010), and Holtug (2001). See also, Adler (2009), and Adler (2011) for similar ideas. Some seeds for this approach seem to have been planted already in Parfit (1995), appendix G, p. 490.

2. Non-identity cases

Non-identity cases are cases where the alternative outcomes do not contain the same people, but they might contain the same number of people. There are many different kinds of non-identity cases, but there are five groups that are especially morally salient (below ‘good/bad life means ‘good/bad for the person leading the life’):

Good lives versus no lives: The mere addition of a number of people with good lives, or no addition at all.

Bad lives versus no lives: The mere addition of a number of people with bad lives, or no addition at all.

Good lives versus bad lives: The mere addition of a number of people with good lives or the mere addition of a number of different people with bad lives.

Good lives versus even better lives. The mere addition of a number of people whose lives are good for them or the mere addition of a number of different people with better lives.

Bad lives versus even worse lives. The mere addition of a number of people with bad lives or the mere addition of a number of different people with worse lives.

There is no general agreement on how to assess these different mere additions. For example, some think that in the case of *Good lives versus no life*, a mere addition of a good life always makes the world better, whereas others deny it. But there is a wide agreement that not every case is a matter of indifference (or incomparability); some of these additions do make the world better or worse. For example, most would agree that in case of *Good lives versus bad lives*, a mere addition of huge number of very good lives (all equally good) makes the world better than a mere addition of a huge number of very bad lives. Similarly, in the case of *Bad lives versus no lives*, most would agree that the mere addition of a huge number of very bad lives is worse than no addition at all.

3. The person-affecting constraint

A common formulation of the person-affecting constraint is this:

The person-affecting constraint (PAC):

If A is better (worse) than B, then A is better (worse) than B for someone.

This formulation calls out for some clarifications. First, it is assumed that PAC is necessarily true and that A and B are *abstract alternatives*, in the sense that they can exist without being instantiated, but they cannot both be instantiated. It is important that A and B are alternatives that cannot both be instantiated, since we are comparing alternatives in which some people exist with alternatives in which these people do not exist. I shall accept these assumptions but not take a stand on the exact ontological nature of the alternatives and the instantiation relation. The alternatives can be seen as abstract *states of affairs* that can exist without *obtaining or being realized*, *properties* that can exist without being *exemplified*, *event-types* that can exist without being *tokened*, or *propositions* that can exist without being *true*. For ease of exposition, I am going to use ‘exemplify’ as a catch-all term for these more specific instantiation relations.

Second, to make sure we have all the information needed to make overall assessments of the alternatives, I shall assume that they are *consistent* and *complete* with respect to what matters for well-being (including enablers and disablers, if there are such things). At times, I shall use ‘S’s existence’ (‘S’s non-existence’) as a convenient way of referring to any abstract complete and consistent alternative that includes (precludes) the existence of S, in the sense that, necessarily, if it were exemplified, S would (not) exist.

Third, I am bracketing issues about non-welfarist values, since I focus on the part of morality that is concerned with well-being, ‘benevolence’, as we might call it. But I also bracket the value of inequality of wellbeing and other welfarist ‘pattern’ values. Without this bracketing PAC would be controversial even in cases where the identity of people is not at stake. For example, if one gives a lot of weight to inequality of wellbeing, making someone worse off need not make things worse overall, since this can be a way to reduce inequality (by ‘levelling down’) without too much of a sacrifice in total wellbeing.

Fourth, there is an obvious problem with this formulation of PAC, if the quantifier ‘for someone’ ranges only over *actual* people. Suppose A and B contain only *non-actual* individuals, i.e., individuals that do not exist in the actual world, and that in A

everyone is extremely happy (and equally so) and in B everyone is extremely unhappy. Then PAC entails that A is not better than B, for there is no actual person for whom A is better than B.

A formulation that avoids this problem is this one:

If A is better (worse) than B, then A is better (worse) than B for someone *who exists in A or in B*.⁵

More precisely, unpacking the disjunctive ‘in’-locution:

If A is better (worse) than B, then either A would better (worse) than B for someone, if A were exemplified, or A would be better (worse) than B for someone, if B were exemplified.

Finally, I shall assume that A is better for S than B if and only if A, S, and B exist and are standing in the relation expressed by ‘__is better for __ than __’. Clear evidence for this is that ‘A is better for S than B’ entails (a) there is *something* that is better for S than B, (b) A is better for S than *something*, and (c) A is better than B for *someone*.

4. The tension between PAC and our intuitive judgements about non-identity cases

Here is an instructive argument for the claim that if the person-affecting constraint is true, then there is no non-identity case of the mere addition kind listed above in which one alternative is better or worse than another and in which. Let A and B be two alternatives in such a case, so

(i) A and B differ with respect to the identity of people, but are equally good for those who exist in both.⁶

Now we add PAC

(ii) If X is better (worse) than Y, then X is better (worse) than Y for someone who exists in X or in Y.

⁵ The need for this reformulation is pointed out in Holtug (2001).

⁶ Alternatively, we can assume, a bit unrealistically, that A and B do not share any individuals.

And two plausible general principles:

(iii) If X is better (worse) than Y for a person who exists in X or in Y, then she would have been better (worse) off in X than in Y (*Better-for entails better-off*).⁷

(iv) No one can be better off (worse off) existing than not existing. (*Well-being entails being*).

which means that

(v) No one in A or in B is better (worse) off in A than in B. (from (i) and (iv))

(vi) A is not better (worse) than B for anyone in A or in B. (from (iii) and (v))

And we get the conclusion

A is not better (worse) than B. (from (ii) and (vi))

Since we chose A and B arbitrarily, this conclusion shows that there is *no* non-identity case of the mere addition kind in which one alternative can be said to be better than another. But this is very worrying, since, as I pointed out in section 2, we would like to say that, at least in some non-identity cases, one alternative is definitely better than another. For example, the mere addition of a huge number of very good lives (all equally good) is definitely better than the mere addition of a huge number of very bad lives.

Recently, there have been attempts to resist this conclusion while holding on to PAC. Some (Roberts, Voorhoeve, and Fleurbaey) reject *Well-being entails being* and claim that non-existence does not preclude being better off (or worse-off).⁸ Others (Adler, Arrhenius, Rabinowicz, Holtug, Johansson) instead reject *Better-for entails better-off* and claim that existence can be better for a person than non-existence even though the person would not be better off existing than not existing. A third option is to deny (i), i.e., deny that there are any non-identity cases, because one thinks that in all the worlds in which a person is not conceived, she still exists as a

⁷ Adler (2009), p. 1503, also states this principle, but he rejects it.

⁸ Roberts (2015) does not defend PAC, but a weaker principle she calls 'the person-based intuition', according to which an outcome A is worse than an outcome B only if A is worse for someone than some alternative outcome Z, where Z need not be identical to B. However, she would have to defend PAC, if it is restricted to cases where A and B are the only available outcomes.

merely possible person, who has wellbeing. I shall argue that none of these ways of blocking the argument works. This leaves PAC itself as the only remaining culprit.

5. Against *Well-being entails being*: Voorhoeve and Fleurbaey

Voorhoeve and Fleurbaey claim that there are non-existent persons.⁹ This claim smacks of incoherence, since it is not easy to find a coherent interpretation of ‘There are persons who do not exist’. After all, to be a person, at least on the most natural interpretation of ‘person’, requires existence. They could drop the reference to persons and just say there are *individuals* who do not exist. But this is not much better, since ‘there are’ seems to have existential import: to say that there are individuals who do not exist is to say that there *exist* individuals who do not exist.

Fleurbaey and Voorhoeve agree with *part* of spirit of the principle *Wellbeing entails being*, for they accept that non-existent individuals have ‘no wellbeing level’. But they nevertheless claim that non-existent individuals can be better-off or worse-off. More exactly, they claim that if S exists in B but not A, it can still be true that if A *were* the case, then A *would* be better for S than B. But note that S is better off in A than in B, if it is both true that A would be better for S than B if A were the case, and B would be worse for S than A if B were the case.¹⁰ How can one accept that one can be better off never existing without assigning well-being levels to the never existing? Their answer:

(...) one can sensibly hold that a particular life can be better for a person than never existing without assigning a level of well-being to never existing. It is sufficient that there is a level of wellbeing, when existing, that is deemed *equivalent* to never existing. (Call this the ‘personal-value indifference’ value of wellbeing; it is often referred to as the ‘neutral level’). Then, we submit, enjoying a greater wellbeing than this level implies that a person’s life is better for her than not existing.¹¹

Even though Voorhoeve and Fleurbaey state that the conclusion is that a person’s life is better for her than not existing, it is clear from the context that they think it also establishes the claim that a person would be *better off* existing and leading the life than not existing.

⁹ Voorhoeve and Fleurbaey (2015), p. 98–100.

¹⁰ Voorhoeve and Fleurbaey (2015), p. 100.

¹¹ Ibid.

What does ‘deemed equivalent’ mean? Deemed equivalent in terms of what? They claim that if a life L has a higher level of well-being than a life L^* and a life L^* is ‘deemed equivalent’ as non-existence, then L is better for the person than non-existence. Surely, for this to work ‘deemed equivalent’ must mean ‘truthfully deemed equivalent *in value* for the person’, or, more succinctly, ‘equally as good for the person as’. But then if L^* has a certain level of wellbeing for a person and L^* is equally as good for her as her non-existence, then her non-existence also has the same level of well-being. In general, if A has a certain level of well-being for S , and A is equally as good for S as B , then B has the same level of well-being for S as A . Indeed, this holds for other comparatives too: if I have a certain height and you are as tall as me, then you also have the same height. If I have a certain weight and you are as heavy as me, then you have the same weight. So, I can’t see how Voorhoeve and Fleurbaey can avoid assigning a level of wellbeing to non-existence.

In any case, their account directly runs into problems with the following principle, which I take to give a defining feature of better off:

Wouldy Better-off

S is better off in A than in B iff the value A would have for S , if A were exemplified, is greater than the value B would have for S , if B were exemplified.¹²

This principle is very plausible, since, as Broome has pointed out, to be better off in A than in B is not just to stand in a relation to A and B ; it involves standing in relations *in* A and *in* B .¹³ But, I would like to add, to have stand in a relation *in* X , where X is a complete and consistent abstract alternative, must be understood *counterfactually*: one would stand in the relation, if X were exemplified. Since nothing can have a value for a non-existent person – a claim Voorhoeve and Fleurbaey themselves come very close to endorsing, since they say that non-existent persons have no level of wellbeing – the above principle rules out their account.¹⁴

¹² As I will explain in section 8, it is not just that A would have value for S , if A were exemplified; it is also true that some part of the world would also have value for S in virtue of exemplifying A .

¹³ Broome (2004).

¹⁴ It is true that it is popular to say things like ‘I would have been better off never existing’, and ‘I would have been better off if I had never been born’, but this must be understood as *hyperbolic idioms*. Taken literally, they are clearly false, if existence is understood in its minimal generic sense of being something, i.e., being identical to something. After all, it is not popular to say ‘I would have been better off not being anything, i.e., not being identical to anything’. In section 7, I shall show that they are false even if we accept that I could have existed as a merely possible person, if I had not been born, or brought into a concrete existence. In any case, we know what is *communicated* by these false idioms: I am really *badly off* existing, which is a claim that does not require any wellbeing comparisons between existing and not existing.

6. Against *Well-being entails being*: Roberts

Roberts accepts that a person can only have properties in situations in which she exists, but she maintains we can still say that a person can have zero level of well-being in situations in which she does not exist. She maintains that:

(...) Nora does not have any properties at all at any alternative at which she does not exist and (...), where Nora has no properties at all, all the properties that she does have – the empty set – add up to zero level of wellbeing.

And this means, she claims, that we can say that the person better off leading a good life than not existing at all. Now, it is not clear what she means by the empty set of properties ‘adding up’ to zero level of wellbeing. Obviously, she cannot mean that to have zero level of wellbeing is the *same* as not having any properties, for I can have zero level of wellbeing and still have properties, for instance, by being unconscious and not having any good or bad things happening to me, or, alternatively, by having good and bad things happening to me when the good things exactly balances the bad things.

A recent argument for her claim that one can be better off existing than not existing is her ‘zero money - zero wellbeing’ analogy. Roberts suggests that to say that a person has zero wellbeing is analogous to saying that she has zero money. Since you can have zero money in a country in which you do not live (China, for instance, in Robert’s case), you can also have zero wellbeing in a world in which you do not exist. The analogy is not convincing.¹⁵ To say that one has zero money in China is to say that one has *no* money in China. By analogy to say that one has zero wellbeing in a world in which one does not exist is to say that one has *no* wellbeing in this world, neither positive, negative, nor neutral wellbeing. But then we cannot say that the person is better off existing than not existing. For, as pointed out in the previous section, ‘better off’ is *wouldy* in the sense that to be better off in A than in B requires that one A would have a certain value for one, if A were realized, and also that B would have a certain value for one, if B were realized.

But perhaps Roberts assumes that to say that S has neutral wellbeing is just to say that it is *not the case* S has overall positive or negative wellbeing. This is one way to understand her claim that Nora’s empty set of properties somehow ‘adds up’ to a zero level of wellbeing. But this is not true. Lots of things lack positive or negative wellbeing, but it is not true to say that they therefore must have a neutral level of well-being. Abstract entities such as number 2, for instance, or certain concrete things, such as my socks, do not have any wellbeing whatsoever.

¹⁵ For a different criticism of Roberts’ analogy, see Johansson (2010), p. 289.

Note that this is not a quibble about terminology. To have neutral wellbeing is evaluatively and normatively significant in a way that lacking any wellbeing whatsoever is not. If a person has neutral wellbeing, we can talk about the *equality* and *inequality* between her wellbeing and others', and we can say that she is *better off* and *worse off* than others. If she has neutral well-being, then if we make her better off, things will be *good* for her, and if we make her worse off, things will be *bad* for her. Furthermore, the fact that she has neutral wellbeing has normative implications. For example, it is then fitting to take a neutral attitude towards her situation. It is also fitting to feel *sympathy* for her, when, through no fault of her, she ends up having only neutral wellbeing in a situation in which she could have had great positive wellbeing. None of these implications hold when a person lacks wellbeing altogether.

It should be noted that in deciding whether it makes sense to say that one is better off existing than not existing it is easy to be misled by locutions such as

There is more value for S in A than in B.

There is more well-being for S in A than in B.

for these sentences can be true because (respectively):

There is some value for S in A but *none* in B, for S does not exist in B.

There is some wellbeing for S in A and *none* in B, for S does not exist in B.

The same phenomenon shows up in other contexts. There is more water in the ocean than in the dry desert, since there is some water in the ocean and none in the dry desert. There is more pain in this world than in a world in which there are no sentient beings, since there is some pain in our world and none in sentience-less world. Or to take an example that is closer to Roberts' own example: I have more money in Sweden than in China, for I have some money in Sweden and none in China. In general, there is a sense of 'more F' that can be used to state that there is more F in X than in Y, when there is some F in X and none in Y.

But 'more wellbeing' and 'more value for', in this sense, should not be conflated with 'being better off'. If A is good for S and S has no wellbeing in B, because S does not exist in B, then, S cannot be better off in A than in B, for this would imply the false claim that B would have a value for S, if B were exemplified.

Roberts' insistence that non-existent persons can have zero level of wellbeing

might thus be based on a conflation between two different readings of ‘zero level of wellbeing’, namely, ‘no wellbeing’ and ‘neutral wellbeing’.¹⁶ Roberts is right to say that there is more wellbeing for me in an alternative A, in which I exist and have a good life, than in an alternative B, in which I do not exist. But this does not show that B would be neutral for me, if B were realized, which is the claim she needs to establish in order to show that I would be better off in A than in B.

7. Against genuine non-identity cases: necessarily existing merely possible persons

A more radical reaction to the argument is to deny that there are any genuine non-identity cases, where the outcomes differ in terms of who exists in them.¹⁷ If you had not been conceived, you would still have existed but only as *a merely possible person*.¹⁸ Furthermore, the daughter you could have had but in fact did not have, still exists as a merely possible person. More generally, if a person exists in a world, as a person, then she exists, as a merely possible person, in all worlds in which she is not conceived. So, there cannot be any non-identity cases, for every person exists necessarily, (but not necessarily as a person).

A merely possible person is something that is in fact not a person but could have been a person.¹⁹ In general, if x is a merely possible F, it does not follow that x is F. Note that a merely possible person is not the same as a *potential* person. A potential person is a *concrete* thing – for example, a fertilized egg – that is not a person but can *develop* into a person. A merely possible person is a non-concrete thing that could have been a person (and concrete). I am not identical to the fertilized egg that developed into me, but I am identical to some merely possible person in a world in which I am not conceived. Of course, that there are necessarily existing merely possible persons in this sense is a very controversial metaphysical assumption, but let this pass and assume for the sake of the argument that there are such beings.

The important question is whether merely possible persons can have wellbeing.

¹⁶ Holtug (2001) seems to make a similar conflation: he equates having nothing good or bad happening to you with having neutral wellbeing.

¹⁷ Williamson (2013), p. 63, suggests that this view could have important ramifications for ethics: if one believes in necessarily existing merely possible persons one cannot simply dismiss as meaningless the claim that it would have been better for a person not to have been born.

¹⁸ This is assumed without arguments in Hare (1988), p. 281.

¹⁹ I am here ignoring David Lewis’ alternative reading of ‘merely possible persons’. On his infamous realist view of possible worlds, presented in Lewis (1986), all possible worlds are concrete, but they are causally and spatiotemporally isolated from each other. To be a merely possible person in a possible world w is to be a concrete person that does not exist in w, but who exists in a *different* possible world. This alternative reading of ‘merely possible person’ will not be of any help in easing the tension between PAC and our judgements about non-identity cases, for someone who is a merely possible person in a possible world w, in this Lewisian sense, does not exist in w and thus cannot have any wellbeing in w.

If they cannot have wellbeing, then we cannot evade the tension between PAC and our considered judgements about some that some purported non-identity cases, for example, that it is better to create a very happy person than to create a very unhappy person. It does not help to be told that the happy person would still have existed as a merely possible person, had she not been conceived, if she then would have lacked any wellbeing whatsoever. If she would have lacked any wellbeing whatsoever, then we can still not say that she would have been better off conceived than not conceived. For, as *Wouldy Better-off* demands, if you are better off conceived than not conceived, then you would have had some wellbeing, if you had not been conceived.

So, the crucial question is 'Is it possible for a merely possible person to have well-being?' Or, more exactly, 'Is it possible that: there is a merely possible person, who has well-being?' We are not interested in the *narrow scope* reading of the question: 'Is there is a merely possible person such that it is possible that she has well-being?', for an affirmative answer to this question does not help. In order to truthfully say that having been unborn *would* have been worse for you than having been born we would need to assign wellbeing to you as an unborn merely possible person. It is not enough to say that you, as a merely possible person, could have had a well-being, because you would have had it, *if* you had been born.

It is very doubtful that you could have had wellbeing, if you had not been conceived. For if you had not been conceived, you would have been a merely possible person, a merely possible animal, a merely possible human, a merely possible philosopher, ... Why should we not continue this list with 'a merely possible bearer of *well-being*? After all, it is very plausible to think that having well-being requires being concrete in some way, for example, having a mind, body, being in space or time, or having causal power.

One could perhaps object that to have well-being it is enough that one has the *capacity* to be concrete in a certain way, having a mind, say.²⁰ Since merely possible people, if they exist, are such that they possibly have minds (they have minds in all worlds in which they are born and develop into mature persons), they also have the capacity to have a mind. So, if you had not been born, you would have been a merely possible person with the capacity to have a mind. And since a capacity to have a mind is sufficient for having well-being, you can be assigned well-being as a merely possible person.

This argument goes wrong in assuming that to have a capacity to have F it is enough that one is such that one possibly has F. This is not true. For example, I am such that I possibly jump to the moon, but I do not have the capacity to do it.²¹ To have a capacity to do or have something requires more than just having the purely

²⁰ This argument is discussed in Bradley (2009), Chapter 3, Section 3.5.

²¹ Johansson (2010) makes this important observation.

modal feature of being such that one possibly does or have it. Arguably, what is required is that the capacity is somehow grounded in features that are not purely modal.

Another argument for the possibility of assigning well-being to merely possible persons would go like this. It is true that to have neutral well-being is not just to lack the properties of having overall positive or negative well-being. To have neutral well-being is a property in its own right. But as a merely possible person you do have the *negative* property of being such that you do not have overall positive or negative well-being, and this property is just the property of having neutral well-being.

This argument overgeneralizes, however. Everything that could not have positive or negative well-being has the property of being such that it has no positive or negative well-being. For instance, the number 2, the null set, and the grain of sand in my pocket all have this property. If the argument were right, we would have to say that these things too have neutral well-being, which of course would be absurd.

One could avoid this objection by revising the definition of neutral wellbeing in the following way: to have neutral well-being is to be a *welfare bearer* (welfare subject) who has the negative property of lacking overall positive and negative wellbeing. Obviously, the number two, the null set, and the grain of the sand in my pocket are not welfare bearers, since it is inconceivable that they have any wellbeing whatsoever.

This maneuver would block the absurd conclusion, but it is important to note that it cannot be used more constructively to show that merely possible people have wellbeing. To ask whether something is a welfare bearer is just to ask whether it fulfils the requirements for having well-being and I have previously shown that it is very doubtful that they do that. Recall that they are not concrete and thus cannot have any of the properties or capacities that requires concreteness.

The prospects look bleak for finding a convincing argument for the conclusion that merely possible people could have well-being levels. They seem too ‘thin’, metaphysically speaking, since they only exemplify trivial or logical properties, such as being self-identical, and purely modal properties not grounded in other properties, such as being possibly F, for any F that the individual exemplifies in a world in which it is concrete.

Apart from these more metaphysical considerations, one could point out the *normatively* absurd consequences of accepting that merely possible persons have wellbeing. This objection has to do with the fact that the notion of well-being is closely tied to reasons to care and feel sympathy. When someone is made worse-off (for no faults of her own) we have reason to care and feel sympathy for her. Consider a possible scenario in which my parents decide not to have a second child and, as a consequence, I am not born. According to the account in question, I am made worse-

off by their decision. In fact, assuming that I would have had a great life, if conceived, I am made *significantly* worse-off. Now, being made significantly worse-off gives people reason to feel sympathy. So, a neutrality account implies that people in this scenario have reason to feel sympathy for me as an unborn merely possible person. But this seems absurd, so the account must be mistaken.

One reply here is to say that we only have reason to feel sympathy for people who are made worse off by being caused to *suffer*. And when my parents decided not to conceive me, they did not cause me to suffer. But this is an all too narrow understanding of sympathy. We can have reason to feel sympathy towards people who do not suffer. For example, we have reason to feel sympathy towards people who are being made unconscious and miss out on good experiences and activities.

Another reply is to say that we have reason to feel sympathy only towards people we *know* about. In the scenario in which I am not born I exist, but since no one knows about me in this scenario no one has reason to feel sympathy for me. But this reply ignores the possibility that we can know about someone just by being able to *refer* to the person. For example, consider a scenario in which my parents decide not to have intercourse, but if they had decided, I would have been conceived. Then they could know about me under the description 'the person who would have been conceived if had had intercourse now'. Of course, this is not a particularly vivid representation of me, but reasons to feel sympathy do not require vivid representations of victims (even though they make it easier to feel sympathy). Just knowing that your closest neighbor was made significantly worse off gives you some reason to feel sympathy even though you know him only under the description 'your closest neighbor'.

All in all, then, there are both strong metaphysical and normative arguments against the possibility of there being merely possible persons with well-being.

8. Against *Better-for entails better-off*: Arrhenius and Rabinowicz

Common to all Scandinavian existentialists is that they think that 'better-for' expresses a relation between abstract states of affairs and individuals. They also accept that if S exists and state of affairs p is good for S and S does not exist in state of affairs q, then S cannot be better off in p than in q; but they insist that in this case p is better for S than q.²² So, they all deny *Better-for entails Better-off*. Since Arrhenius & Rabinowicz have developed their version of Scandinavian existen-

²² Adler (2009), p. 1505, and Adler (2011), p. 220, makes a similar point as a defence of a preference-based account of wellbeing. He states that even though it does not make sense to say that you could have been worse off not existing, it does make sense to say that your non-existence is worse for you than your existence in virtue of the fact that you prefer your existence to your non-existence.

tialism further than others, I shall focus on them in the following, but my criticisms apply to the whole camp.

Arrhenius & Rabinowicz point out that if *S* exists, then there are no missing relata.²³ We have the two abstract states of affairs *p* and *q*, and also the individual *S*. They also point out that it is not true in general that if a relation *R* holds over abstract states of affairs *p*, *q*, and person *S*, then *R* would hold over *p*, *q*, and *S* even if *p* were to obtain. For example, it does not hold if the relation is preference and the states of affairs are your non-existence and your existence and the person is you. So, the sheer logic of relations cannot be used to establish that if *p* is better for *S* than *q*, then *p* would have value for *S* (comparative or non-comparative), if *p* were to obtain.

I agree that the sheer logic of relations is not helpful here. But I think there is a strong argument against their view. Before I present the argument, we need to consider what it means, more generally, to say that an exemplifiable abstract entity has value. Let us first consider abstract *properties*. What do we mean when we say that properties such as courage and benevolence have value? At least the following, I maintain: If courage is good, then it is possible for people to be good in virtue of exemplifying courage. If benevolence is good, then it is possible for people to be good in virtue of exemplifying benevolence. The same holds for all other virtue properties. The slogan is: a good character trait could ‘rate a person a plus’.

Similar things can be said about functional goodness. If sharpness is a good property for knives, then it is possible for knives to be good in virtue of exemplifying sharpness. In short, sharpness could ‘rate a knife a plus’. If empathy is a good property for a counsellor, then it is possible for counsellors to be good in virtue of exemplifying empathy. In slogan form, empathy could ‘rate a counsellor a plus.’

As Sven Danielsson, drawing on Chisholm, has suggested, this idea can be generalized to the value of *states of affairs*.²⁴ If a state of affairs *p* is good, then it is possible for something (‘the universe’, Chisholm labels it) to be good in virtue of exemplifying *p*. For example, if the state of affairs my feeling pleasure is intrinsically good, then it is possible that something is intrinsically good, at least to some extent, by exemplifying this state of affairs. As Chisholm puts it, the slogan is: A good state of affairs could ‘rate the universe a plus’.²⁵

Finally, applied to well-being the idea is that if a state of affairs *p* is good for you, then it is possible for something to be good for you in virtue of exemplifying *p*. For

²³ Holtug (2001), pp. 139-140, makes the same point.

²⁴ Chisholm (1966), Danielsson, unpublished manuscript.

²⁵ In fact, I think this idea can be generalized even further to the normative status of actions. If an abstract act-type, such as lying or killing, is wrong, then it is possible that some act-token of these types is *pro tanto* wrong in virtue of exemplifying the type in question. Indeed, I think the idea can be generalized to some *non-moral* cases too. For example, it seems natural to think that if eating vegetables is healthy, then it is possible that someone is made healthy (to some extent) by exemplifying the act-type eating vegetables.

example, if the state of affairs of my feeling pleasure is good for me, then it is possible that something is good, at least to some extent, for me in virtue of exemplifying this state of affairs. The slogan is: states of affairs that are good for you could rate something a plus for you. They are possible good-for-you makers.

It may not be perfectly obvious why this idea is a natural generalization of the idea that good properties are possible good-makers. But this generalization is difficult to resist once it is acknowledged that abstract states of affairs are similar to ordinary properties. States of affair are most plausibly seen as *ways* things could be. After all, they are states *of* something, namely, ‘affairs’. If states of affairs are ways things could be, then possible worlds are maximal consistent states of affairs, i.e., *maximal* ways things could be. There is only one world, ‘the totality of what exists’, ‘you and all your surroundings’, and what we call possible worlds are more adequately called possible world-*states*, i.e., maximal ways the world could be.²⁶ A state of affairs *p* is exemplified just in case the world exemplifies a world-*state* that entails *p*. A state of affairs *p* obtains *in* a possible world-state *w* just in case *w* entails *p*.²⁷

So far, we have considered *goodness* of abstracta that can be exemplified. But the same idea applies to badness, neutrality, or any other kind of value of such abstracta²⁸ The general principle, *Possible Value-making*, is that, for all exemplifiable abstracta *X*, if *X* has a certain value, then it is possible that something has this value in virtue of exemplifying *X*. For example, if *X* is good, then it is possible that something is good in virtue of exemplifying *X*; if it is bad, then it is possible that something is bad in virtue of exemplifying *X*, and if it is neutral, then it is possible that something is neutral in virtue of exemplifying *X*. In slogan form: an *X* that has value could rate some *X*-exemplifier this value. So, a good *X* could rate some *X*-exemplifier a plus, a bad *X* could rate some *X*-exemplifier a minus, and a neutral *X* could rate some *X*-exemplifier a zero.

The exemplifier of an abstract *X* is (some part) of *the* world, in the sense of ‘the totality of things’, or ‘you and all your surroundings’. What the exemplifier consists in depends on the nature of *X*. The exemplifier is a token event, if *X* is an event-type;

²⁶ I take no stand on what *the world*, ‘you and all your surroundings’, consists of. In particular, I do not assume that it only consists of concrete individuals.

²⁷ States of affairs behave like properties even if one follows David Lewis and takes the notion of a concrete possible world as primitive and identify abstract states of affairs with sets of concrete possible worlds, for then states of affairs can be seen as extensional properties of concrete possible worlds. A state of affairs is exemplified just in case the actual concrete world is a member of *p* (i.e., the set of all the *p*-worlds). A state of affairs *p* is exemplified by a concrete possible world *w* just in case *w* is a member of *p*.

²⁸ Examples of other kinds of the value would be the indeterminate values of disjunctions that consist of disjuncts with different absolute values (e.g., being happy or being unhappy, being happy or being indifferent), and what Gustafsson (2016), p. 855, calls the ‘blank’ values - whatever has such a value can be worse than what is good and better than what is bad but nevertheless fail to be equally as good as what is neutral. Carlson (2016) defends a similar idea.

a particular, if it is a property; a truth-maker, if it is a proposition; and a realization of a state of affairs, if it is a state of affairs, where the realization in turn consists of 'states' i.e., properties and relations, and parts of the world (events, processes, particulars, or individuals) that exemplify these states.

This means that I am *not* talking about value-making in terms of *entailment*: one *abstract* alternative X making another *abstract* alternative Y have a certain value by being part of Y, or being entailed by X. Here are some examples of this kind of value-making:

The property of being generous makes the complex property being generous and brave better by being part of it.

The property of being sharp makes the complex property of being sharp and easy to clean better by being part of it.

The state of affairs that John feels pleasure makes the complex state of affairs that John feels pleasure and Jane feels pleasure better by being entailed by it.

This kind of value-making holds between abstract alternatives and must be distinguished from the value-making I have in mind: value-making by way of *exemplification*, which holds between abstract alternatives and their exemplifiers. Here are some examples: being generous makes a *person* better, being sharp makes a *knife* better, and that John feels pleasure makes *the world* better (by making the concrete situation, in which John is feeling happiness, better). There are thus two important analogies between valuable properties and valuable states of affairs. Both are possible value-makers by way of entailment as well as possible value-makers by way of exemplification. But it is only the latter kind of value-making that is invoked in the principle *Possible Value-Making*.

To avoid misunderstandings, a few further clarifications are in place.

(i) The explanation expressed by 'something has value in virtue of exemplifying X' is *inclusive* in the sense that enablers and disablers (if there are such things) are part of the explanation. There are less inclusive notions of 'in virtue of' which would leave out the enablers and disablers and only list the factors that are enabled and not disabled to make things good, bad, or neutral.

(ii) As it stands, the principle *Possible Value-Making* is best suited to express claims about *final or intrinsic value*, since if X has this kind of value it seems clear

that it is in virtue of exemplifying X that things can have this value.²⁹

(iii) The account does not say that only *concrete* things can be good in virtue of exemplifying abstract features. What exemplifies a good abstract feature can itself be abstract. So, for instance, a musical composition, assuming it to be abstract, can exemplify elegance, simplicity, and balance and be good (beautiful) in virtue of exemplifying these good abstract features.

(iv) Perhaps it is true that to say that properties and states of affairs are good is *just* to say that they are possible good-makers. I am inclined to think that this is true, but I need not take a stand on this issue here. Perhaps what is a possible good-maker must also itself be good (and not just a good-maker) – what contributes to the goodness of things must itself have some goodness to contribute, as Dancy puts it.

(v) The possibility invoked in *Possible Value-Making* can be given different interpretations. Metaphysical possibility is one option. But some may claim that, on this interpretation, the account rules out too much, for isn't the state of affairs of *John's being a very happy cow* good even though it is metaphysically impossible that John is a very happy cow? Or, could we not say that the state of affairs of *there being a god who enjoys her creation* is good even if it is metaphysically impossible that there is a god? I think that we should not say that these states of affairs *are* good, only that they *would* have been good, had things been radically different (indeed, so different that essential truths would have been different). However, for my argument I only need weaker versions of the principle that invokes *logical* possibility, in the sense of what is compatible with the truths of logic. So understood, the *Possible Value-Making* will not rule out that the above-mentioned states of affairs are good, since it is metaphysically but not logically impossible that John is a happy cow and that there is a god.

If we accept this understanding of the value of abstract exemplifiable alternatives, we have an argument against the idea that existence can be better for you than non-existence. For all A, B, and S (all premises being necessarily true):

²⁹ As regards the *extrinsic* value of X, which is defined in terms of the final or intrinsic value of features other than X that would be exemplified if X were exemplified, it seems at least misleading to say that things can be good in virtue of (just) exemplifying X. Rather, here things can have value in virtue of exemplifying all the valuable features that were exemplified as a result of X being exemplified. To make room for this kind of value, we could simply weaken the principles further so that they say that if X is good (bad/neutral), then it is possible that something is good (bad/neutral) *and X is exemplified*. No matter whether X or its counterfactual dependents make things good, X itself must be exemplified. Similar considerations apply to states of affairs with *non-basic* intrinsic value. For example, suppose a certain exclusive disjunction of good states of affairs is itself good (e.g., your being happy to degree 5 or your being happy to degree 10), and that it is good in virtue of the good disjuncts, then one may want to claim that when this state of affairs is exemplified things are made good, not in virtue of the whole disjunction, but in virtue of the disjunct in virtue of which the disjunction is exemplified.

(1) If A is better (worse) for S than B, then the value of A for S is greater (less) than the value of B for S. (*Better-for entails value-for*)

(2) If A has a certain value for S, then it is possible that something has this value for S in virtue of exemplifying A. (*Possible Value-Making*)

(3) If something has a certain value for S in virtue of exemplifying A, then A is exemplified and S exists. (*Factivity*)

So

(4) If A has a certain value for S, then it is possible that A is exemplified and S exists. (from (2) and (3))

So

(5) If A is better (worse) for S than B, then it is possible that A is exemplified and S exists and it is possible that B is exemplified and S exists. (from (1) and (4))

So

(6) If the state of affairs of S's existence is better (worse) for S than the state of affairs of S's non-existence, then it is possible that S's existence is exemplified and S exists and it is possible that S's non-existence is exemplified and S exists. (from (5))

(7) It is not possible that S's non-existence is exemplified and S exists. (logical truth)

So

S's existence is not better (worse) for S than S's non-existence. (from (6) and (7))

Premise (1), *Better-for entails value-for*, states a very plausible principle. Indeed, something similar holds for all comparatives. If I am taller than you, then my height is greater than your height; if I am heavier than you, then my weight is greater than

your height. In general, if x is F_{er} than y (and ‘F_{er} than’ is a comparative), the F_{ness} of x is greater than the F_{ness} of y.³⁰

Premise (2) is true since it is an application of the general principle *Possible Value-Making* to ‘value for’. Remember that we are here talking about value-making by way of exemplification, which holds between abstract alternatives and their *exemplifiers*, not value-making by way of entailments, which holds between abstract alternatives when one valuable alternative is part of or entailed by another.

Premise (3), *Factivity*, is true, since ‘value for S’ entails the existence of S, and ‘in virtue of’ is factive: if x is F in virtue of x being G, then x is F and x is G and thus x exists.

Premise (7) is just a logical truth: necessarily, if S’s non-existence is exemplified, then S does not exist.

Note that this argument, if sound, enables us to restate the challenge for PAC without relying on *Better-for-Better off* (though I do believe that this principle is also true). Its conclusion together with PAC entails the problematic conclusion that in no non-identity case (of the mere addition form) can one alternative be better or worse than another.

9. Better-for does not entail better-off: guardian angels

At this point the opponents could say that even though some states of affairs, such as my non-existence, are not possible value makers, they can still merit certain *attitudinal responses* by existing people.³¹ Drawing on this idea, Arrhenius & Rabinowicz present the following argument for the claim that a person’s existence can be better for her than her non-existence.³² We assume throughout that we are talking about final values for S and corresponding final preferences.

(1) If A is good for S and S does not exist in B, then one ought to prefer A to B, for S’s sake.

³⁰ That the F_{ness} of x is greater than the F_{ness} of y does not mean that F_{ness} must always come in precise or determinate levels or degrees, for there can be some indeterminacy about F_{ness}. In that case the F_{ness} of something is better represented by a *range* of precise levels of F_{ness}. In order to compare the F_{ness} of x to that of y we would then need to compare the range of precise levels associated with x with the range associated with y. One simple idea would be to say that one range is greater than another if all precise levels in the former range is greater than those in the latter, but there are alternative ways of making sense of this comparison.

³¹ Arrhenius & Rabinowicz (2013).

³² A similar argument is presented in Bradley (2014).

(2) If one ought to prefer A to B, for S's sake, then A is better for S than B.

So,

(3) If A is good for S and S does not exist in B, then A is better for S than B.

Since S cannot be better off in A than in B, because S does not exist in B, the argument, if sound, also provides a counterexample to *Better-for entails Better-off*.

'for S's sake' can mean different things.³³ Sometimes

(T1) 'for S's sake' means 'insofar as one *respects* S'.

But then (2) is obviously false, since what is good or better for someone can come apart from what would show her most respect. It can be true that I ought to respect your decision to sacrifice your life for others, and thus that I ought to prefer that you do it, out of respect to you, even though not sacrificing yourself would make you better off.

Arrhenius & Rabinowicz seem inclined to think that 'for S's sake' means 'insofar as one cares about S'.³⁴ But then 'for the sake of S' can mean different things depending on *how* and *why* one cares about S. Is there a uniform interpretation of 'for S's sake' that makes the argument convincing?

(T2) 'for S's sake' means 'insofar as one cares about what is *better* for S'.

This reading would simply be question-begging, since, then, (1) on its own entails that 'If A is good for S and S does not exist in B, then A is better for S than B', for 'ought to prefer insofar as one cares about what is better for S' entails 'better for S'.

(T3) 'for S's sake' means 'insofar as one cares about what is *impartially* better in terms of S's wellbeing'.

I would agree that (1) is true, on this reading, since I think that A is in one respect impartially better than B in virtue of the fact that A is good for S and B is not. However, to insist that this shows that A is *better* for S than B, and thus to assume

³³ I am not suggesting that 'for S's sake of' shows the same ambiguity as 'bank'. It is more plausible to think it has a unified context-invariant meaning that determines what it expresses in a certain context. The standing meaning would then be something like 'on the grounds of some contextually salient purpose that involves S being favoured in some sense'.

³⁴ Arrhenius & Rabinowicz (2015), footnote 22, refers approvingly to a quote from Darwall (2002), in which 'should desire A for S's sake' is equated with 'should desire A insofar as one cares about S'.

that (2) is true, again, would be question-begging, since the issue is exactly whether A is better for S than the S-less B.

(T4) ‘for S’s sake’ means ‘insofar one cares about what contains more wellbeing for S, or more value for S’.

This would make (1) true, since, as I pointed out in my discussion of Roberts’ views, if A contains some wellbeing for S and B none, then, in one sense of ‘more wellbeing’, A contains more of S’s wellbeing than B does, and it seems reasonable to say that one ought, for S’s sake, prefer more wellbeing for S. But, again, it seems question-begging to assume that from the fact that there is more wellbeing for S in A than in B, just because there is some in A and *none* in B, it follows that A is better for S than B. As pointed out earlier, to say that there is no wellbeing for S in B is not the same as saying that B has neutral value for S.

None of these three interpretations gives us a non-question begging uniform reading of ‘for S’s sake’ that would make both (1) and (2) clearly true. That the argument still may seem intuitively attractive can be explained by a shift in our understanding of ‘for S’s sake’. We find (1) true because we assume one interpretation and (2) true because we unwittingly assume another.

A last-ditch attempt to save the argument would be to just insist that there must be an *indefinable* sense of ‘for sake of’ that will fit the bill and make all premises true. But why should we believe that? We need some independent reason to believe in such a primitive sense of ‘for the sake of’. Furthermore, even if we concede that there is such a primitive sense, (2) is still highly contestable. Here are some possible counterexamples that need to be addressed, at least if ‘prefer’ is understood as ‘contemplate with greater pleasure (or some other positive contemplative positive attitude)’:

(a) One ought to prefer, for your sake, your eating disgusting mud to your not doing it, since the evil demon will kill you, if you do not eat the mud, but your eating disgusting mud is not (finally) better for you.³⁵

(b) One ought to prefer, for your sake, your life’s being good for you to your life’s being bad for you, but your life’s being good for you is not better for you than your

³⁵ This is an example of the so-called ‘wrong kind of reason’-objection, which has attracted a lot of attention recently. For a recent paper that has shaped much of this debate, see Rabinowicz & Rønnow-Rasmussen (2004).

life's being bad for you. The value for you of your life's being good for you is not greater than the value for you of your life's being bad for you, since if *evaluative* states of affairs, such as your life's being good for you, themselves have value for you, an infinite regress of values for you looms: A has value for you, (A's having value) has value for you, ((A's having value) having value for you) has value for you...

(c) One ought to prefer, for your sake, your being a happy cow to your being an unhappy cow, even though your being a happy cow is not better for you, since it is *metaphysically impossible* for you to be a cow (assuming you are a human). Of course, if you *had* been a cow, then it *would* have been better for you to be a happy cow.

(d) One ought to prefer, for your sake, your being extremely happy and the only barber who shaves all and only those who do not shave themselves to your being unhappy, but the former state of affairs is not better for you since it is *logically impossible*.

10. Concluding remarks

None of the attempts to ease the tension between PAC and our considered judgments seems promising. A more plausible way to avoid the counterintuitive conclusion that no outcome is better than another in non-identity cases is simply to abandon PAC. I would like to end on a more conciliatory note, however. Arrhenius & Rabinowicz and I agree that a person is better off (or worse off) in one alternative than in another only if she exists in both alternatives. So, we all agree that a person cannot be affected for better or worse by being created. And we also think that an alternative can be better without anyone being better off. Unlike Fleurbaey, Voorhoeve and Roberts, Arrhenius & Rabinowicz agree with me that this principle should be rejected:

If A is better than B, then some person (in A or in B) is better off in A than in B.

So, it is clear then that neither Arrhenius & Rabinowicz, nor I, accept a *strict* person-affecting constraint, according to which what is better must have *affected* some individual. Instead, we accept:

Extended person-affecting principle (EPAC):

If A is better than B, then

(a) some person (who exists in both A and B) is better off in A than in B, or

(b) A is good for a person who exists in A but not in B, or

(c) B is bad for a person who exists in B but not A.³⁶

This is still a person-affecting constraint, but in less strict sense: what is better must make a positive difference in individual well-being facts. After all, to create someone who will be well off is also a way of affecting individual wellbeing, since it is a way to make a positive difference in individual wellbeing facts, even though it does not make anyone better off. It is positive difference in wellbeing facts, since what is added is something that is good for people.³⁷ Similarly, to create someone who will be badly off is also a way of making a negative difference in individual wellbeing facts, even though it does not to make anyone worse off.³⁸ It is negative, since what is added is something that is bad for people.³⁹

The difference between Arrhenius & Rabinowicz and me is that they want to add that in case (b) it is true in A, where S exists, that A is better for S than B, and that in case (c) it is true in B, where S exists, that A is better for S than B. I strongly doubt that it is really worth violating reasonable principles for value-for, such as *Possible Value-Making*, just to be able to maintain a *comparative* form of a person-affecting constraint - 'better' entails '*better* for someone', especially when it is already agreed that non-identity cases show that an alternative can be better without affecting anyone for the better.

³⁶ More precisely: If A is better than B, then

(a) either if A were the case, then it would be the case that someone is better off in A than in B, (or, equivalently, if B were the case, then it would be the case that someone is better off in A than in B),

(b) If A were the case, then it would be the case that A is good for someone who would not exist if B were the case,

(c) If B were the case, then it would be the case that B is bad for someone who would not exist, if A were the case.

³⁷ This could be seen as a *non-comparative* benefit, as McMahan (2013) has recently argued.

³⁸ For more on different forms of person-affecting restrictions, see Arrhenius (2009).

³⁹ This could be seen as a *non-comparative* harm, as McMahan (2013) has recently argued.

References

- Adler, M. D. *Well-Being and Fair Distribution: Beyond Cost-Benefit Analysis*, Oxford University Press, 2011.
- Adler, M. D. "Future Generations: A Prioritarian View", *The George Washington Law Review*, Vol. 77, No. 5/6, September 2009.
- Arrhenius, G. "Can the Person Affecting Restriction Solve the Problems in Population Ethics?" in M. A. Roberts and D. Wasserman (eds.), *Harming Future Persons*. Aldershot: Ashgate, 2009, pp. 289–314.
- Arrhenius, G. "The Affirmative Answer to the Existential Question and the Person Affecting Restriction." in I. Hirose and A. Reisner (eds.), *Weighing and Reasoning*, Oxford University Press, 2014.
- Arrhenius, G. and W. Rabinowicz. "Better to Be Than Not to Be?" in H. Joas and B. Klein (eds.), *The Benefit of Broad Horizons. Intellectual and Institutional Preconditions for a Global Social Science*, *Festschrift for Björn Wittrock*, Brill: Leiden, 2010, 399–421.
- Arrhenius, G. and Rabinowicz, W. 'The Value of Existence' in Hirose, I. and J. Olson (eds.), *Oxford Handbook of Value Theory*, 2015.
- Bradley, B. 'Asymmetries in benefitting, harming, and creating', *Journal of Ethics*, 2014.
- Bradley, B. *Wellbeing and Death*, Oxford: Oxford University Press, 2009.
- Broome, J. *Ethics out of Economics*, Cambridge: Cambridge University Press, 1999.
- Broome, J. *Weighing Lives*, Oxford: Oxford University Press, 2004.
- Carlson, E. "Good' in terms of 'Better'", *Noûs* 50:1, 2016, pp. 213–223.
- Chisholm, R. M. and Sosa, E. "On the Logic of 'Intrinsically Better'", *American Philosophical Quarterly*, 3, 1966, pp. 244–249.
- Darwall, S. *Welfare and Rational Care*. Princeton, NJ: Princeton University Press, 2002.
- Fleurbaey, M. and A. Voorhoeve, "On the Social and Personal Value of Existence." In I. Hirose and A. Reisner (eds.), *Weighing and Reasoning: A Festschrift for John Broome*. Oxford University Press. 2014.
- Glover, J. (1977). *Causing Death and Saving Lives*. New York: Penguin.

- Gustafsson, J. E. "Neither 'Good' in Terms of 'Better' nor 'Better' in Terms of 'Good'", *Noûs*, online prepublication, DOI: 10.1111/nous.12038, 2013.
- Gustafsson, J. E. 'Still Not "Good" in Terms of "Better"' *Noûs*, 50 (4), 2016, pp. 854–864.
- Hare, R. "Possible People." *Bioethics* 2, 1988, pp. 279–93.
- Hirose, I. and Reisner, A. (eds.) (2015) *Weighing and Reasoning: Themes from the Philosophy of John Broome*, Oxford: Oxford University Press.
- Holtug, N. "Person-Affecting Moralities." In J. Ryberg and T. Tännsjö (eds.), *The Repugnant Conclusion: Essays on Population Ethics*. Dordrecht: Kluwer, 2004, pp. 129–61.
- Holtug, N. "In Defence of the Slogan." In W. Rabinowicz (ed.), *Preference and Value: Preferentialism in Ethics*. Studies in Philosophy, vol. 1. Lund: Department of Philosophy, Lund University, 1996, 64–89.
- Holtug, N. "On the Value of Coming into Existence." *Journal of Ethics* 5, 2001, pp. 361–84.
- Johansson, J. "Being and Betterness." *Utilitas* 22, 2010, pp. 285–302.
- Lewis, D. *On the Plurality of Worlds*, Oxford: Blackwell Publishers, 1986.
- McMahan, J. 'Causing people to exist and saving people's lives', *Journal of Ethics* 17, 2013.
- Narveson, J. "Utilitarianism and New Generations." *Mind* 76 (January), 1967, pp. 62–72.
- Parfit, D. *Reasons and Persons*. Oxford: Clarendon Press, 1995 (1984).
- Rabinowicz W. and Rønnow-Rasmussen T. 'The Strike of The Demon. On Fitting Pro-Attitudes and Value', *Ethics* 114, April, 2004, pp. 391–423.
- Roberts, M. A. *Child versus Childmaker: Future Persons and Present Duties in Ethics and the Law*. Lanham, MD: Rowman and Littlefield, 1998.
- Roberts, M. A. "Can It Ever Be Better Never to Have Existed at All? Person-Based Consequentialism and a New Repugnant Conclusion." *Journal of Applied Philosophy* 20 (2), 2003, pp. 159–85.
- Temkin, L. S. *Inequality*. Oxford: Oxford University Press, 1993a.

Temkin, L. S. "Harmful Goods, Harmless Bads." In R. G. Frey and C. W. Morris (eds.), *Value, Welfare, and Morality*. Cambridge: Cambridge University Press, 291–324. 1993b.

Williamson, T. *Modal Logic as Metaphysics*, Oxford University Press, 2013.

Zimmerman, M. J. *The Nature of Intrinsic Value*, Lanham, Md.: Rowman & Littlefield, 2001.